Lariat: A Visual Analytics Tool for Social Media Researchers to Explore Twitter Datasets

Nan-Chen Chen, Michael Brooks, Rafal Kocielnik, Sungsoo (Ray) Hong,
Jeff Smith, Sanny Lin, Zening Qu, Cecilia Aragon
Department of Human Centered Design & Engineering, University of Washington, Seattle, WA, USA
[nanchen, mjbrooks, rkoc, rayhong, jasmba15, sanny, zqu, aragon]@uw.edu

Abstract

Online social data is potentially a rich source of insight into human behavior, but the sheer size of these datasets requires specialized tools to facilitate social media research. Visual analytics tools are one promising approach, but calls have been made for more in-depth studies in specific application domains to contribute to the design of such tools. We conducted a formative study to better understand the needs of social media researchers, and created Lariat, a visual analytics tool that facilitates exploratory data analysis through integrated grouping and visualization of social media data. The design of Lariat was informed by the results of the formative study and sensemaking theory, both indicating that the exploratory processes require search, comparison, verification, and iterative refinement. Based on our results and the evaluation of Lariat, we identify a number of design implications for future visual analytics tools in this domain.

1. Introduction

Increasingly widespread use of social media and online communication platforms produces vast amounts of data. This data is potentially a rich source of insight about human behavior, and it is drawing the attention of social scientists from many fields, including communication [2], organizational systems [8], information and computer science [9, 27], and sociology [30]. Unfortunately, datasets obtained from social media platforms are challenging to work with. They come with a dazzling assortment of metadata, but often lack key information that social scientists desire [26]; they vary unpredictably over time and geography; they often have diverse and unexpected content; and they tend to be large, messy, and incomplete. Because social media data is the result of informal human communication, they are partially structured, contextual, and layered with meaning.

As a result, social media datasets often push the limits of traditional research methods, and researchers find themselves exploring novel and hybrid methods that would better suit this new type of data. Because social media data is large and complex, studying it requires laborious human judgment, consideration of context, and nuanced slicing and cleaning; many of the new methods being explored combine qualitative and quantitative techniques [10, 16]. Regardless of the specific approach taken, the complexity and diversity of social media datasets warrant that care be taken in exploratory data analysis. Researchers must take the time to understand what their datasets contain, and what the contents of these datasets mean in the context of their research.

Because of the above challenges, social media data presents an interesting opportunity for the field of visual analytics [3, 7, 11, 17], which deals with the integration of visual and computational techniques for deriving insight from complex data. To build visual analytics tools that fit into users' workflow, researchers have called for more naturalistic studies conducted in specific application domains, to ground and direct ongoing visual analytics research, design, and evaluation [4, 19]. Toward this end, we conducted an interview-based formative study with social scientists to understand their current needs, practices, and constraints in working with social media data. With input from users, we developed Lariat, an exploratory visual analytics tool for social media datasets which supports iterative category construction and visualization in order to help researchers discover and refine salient categories of content, while also comparing and analyzing categories with visualizations and statistical summaries. Users can explicitly view the original text in addition to the visual comparison, which is critical to researchers using both qualitative and quantitative methods. Furthermore, we incorporate sensemaking theory into Lariat's design where the tool supports iterative search, comparison, verification, and refinement for exploratory processes.

We contribute our domain-specific exploration of challenges and opportunities for visual analytics technology. We offer design implications for visual analytics tools that support social scientists in exploring social media data, including the particular challenges of unifying textual and quantitative aspects of this data, and supporting a mix of qualitative and quantitative analytical techniques. Finally, we contribute the Lariat system which incorporates social

URI: http://hdl.handle.net/10125/41382
ISBN: 978-0-9981331-0-2

media researchers' needs and sensemaking theory to support exploratory processes, and we also conducted a qualitative evaluation to develop a set of design recommendations for exploratory data analysis tools for social media researchers.

2. Background and related work

In this section, we discuss prior work on visual analytics for social media data. We also briefly review exploratory data analysis and sensemaking as theoretical background for our design approach.

2.1. Visual analytics for social media data

Visual analytics researchers have created systems for exploration, analysis, and monitoring of Twitter and social media datasets. For journalists analyzing posts during events, the *Vox Civitas* [7] and twitInfo [17] systems use temporal visualizations and sentiment analysis. For real-time monitoring, the "Visual Backchannel" system presents a stream graph of Twitter topics over time, as well as relevant Twitter usernames, tweets, and photos [6]. Mazumdar et al. developed an approach that displays social media streams in context in order to support analysis by non-technical emergency responders [18], and Hubman-Haidvogel et al. describe a dynamic topography technique that shows how social media topics change over time [11]. Brooks et al. investigated collaborative visual analytics for researchers working with Twitter data [3]. In designing Lariat, we focused on enabling users to approach their datasets from various dimensions, and to let them quickly synthesize groups of data points and make comparisons between groups across these dimensions.

In addition to specific visual analytics tools mentioned above, social scientists also use general-purpose tools, such as Excel, Google Spreadsheets, Tableau, and Gephi, to create visualizations and analyze social media data. Previous work has focused on extending such general purpose tools. Some other tools made visualization a feature of either social media data collection (e.g., Socialpeeks [1]) or qualitative coding (e.g., DiscoverText [25]).

Although many tools exist to support social media analyses through visualization, we wish to call attention to understanding and categorizing users' needs and challenges in this domain. Previous work also suggested that visual analytics needs more naturalistic studies of specific domains, with an eye towards design implications [19]. Several studies have focused on intelligence analysis (IA), uncovering implications for design such as using large screens and affordances from physical media [5], as well as theoretical understanding of the structure of IA work [13]. Similar naturalistic studies have been performed in other domains, e.g. building design [28], automotive engineering [24], and enterprise data analytics [12]. In keeping with the goals of these studies, we also treated Lariat as a means to gain insight on how users interact with visual analytics tools to familiarize themselves with and explore their datasets. We suggest further development of solid design guidelines for building visual analytics tools for the purpose of social media research.

2.2. Exploratory data analysis and sensemaking theory

The field of visual analytics is historically rooted in exploratory data analysis [14], and visual analytics tools are typically designed for exploratory use [20]. In this paper, we focus on the challenges to effective exploration of social media data, as we will discuss in detail later in both formative study results and the discussion section. Exploratory data analysis (EDA) is an open-ended process of discovering the broad structure and characteristics of data, often through iteratively examining summary statistics and visualizations of the data [29]. The purpose of EDA is to discover unexpected hypotheses and findings, often called *insight* in the visual analytics literature [22].

EDA is a complex activity, but sensemaking theory provides a model for understanding how an analyst builds knowledge while exploring data. Sensemaking involves an iterative process of finding schemas, seeking information, comparing schemas and data, and modifying schemas [23]. Pirolli and Card have evaluated sensemaking as a framework to inform visual analytics system design [21]. Klein et al.'s Data-Frame theory provides an alternative perspective on sensemaking based on construction and modification of frames [15]. Kang and Stasko used Data-Frame theory in a longitudinal study of intelligence analysis practices to inform visual analytics system design [13]. Both of these flavors of sensemaking theory emphasize a tightly integrated, interdependent process where the analyst rapidly switches between analysis and synthesis, deduction and induction, brainstorming and refinement. We draw on this theoretical understanding of sensemaking and EDA. While most exploratory tools for tweets may explicitly or implicitly encode the sense-making process, the critical difference with Lariat is to emphasize iterative synthesization and comparison in both textual and visual aspects. The design choice of devoting a large portion of the interface to text was also informed by the social scientists in the formative study. To them, reading tweet text itself is an important part of their workflow. Thus, Lariat leverages sensemaking theory in a domain-specific way, which makes it stand out from other tools.

3. Lariat system design

To design a visual analytics tool for social scientists, we started from a formative study with users. This initial study led us to focus on exploratory data analysis (EDA), and we drew from sensemaking theory to enable keyword-based grouping of tweets as a way to quickly synthesize analysis units and make comparisons. More details are provided in the following subsections.

3.1. The formative study with social scientists

To develop a better understanding of the challenges that our target users face, we conducted semi-structured interviews with four social scientists with at least two years of experience studying social media datasets. All had published at least one top conference paper or journal paper. We sought informants at a large university with an active online social data research community, focusing on researchers with an interest in analyzing the text-based content of their datasets. Participants were asked to focus on a specific recent research project, and tell us about project phases, collaborators, and data. We then asked about challenges they encountered, and more specific questions about their use of visualization tools and the analytical techniques they incorporated. For a detailed analysis of these and other interviews with social scientists studying social media data, see [*omitted for anonymity*].

3.1.1. Datasets. All of our interviewees primarily collected data from Twitter, though other sources were used as well. Interviewees described datasets ranging from a few thousand to millions of tweets, usually collected with a set of filter keywords over a specific time range. According to interviewees, social media datasets are complex, and involve hundreds of metadata fields, which are necessary for analysis: *"Preservation of all of the metadata was actually extremely important and most of the tools were basically like, well which two pieces are important to you and I'll keep those two pieces and then they throw everything out."* However, not all of the relevant information is available in the metadata; interviewees said they also spend a lot of time looking for supplementary contextual information.

3.1.2. Exploratory analysis. According to our interviewees, the complexity of information in social media datasets makes EDA very important, but also challenging: *"Knowing how to approach it [the data] and how to tackle it, how to break it up into different chunks wasn't necessarily clear."* Multiple strategies are used to get a sense of the data: *"It's hard to browse [...] we have all these strategies of… let's look at time series data, let's read through some of the tweets."* Visualization played a role here. One interviewee reported bringing visualizations of the dataset to group meetings to facilitate collaborative exploration: *"Once we had identified these other interesting trends or points of interest, sometimes we'd use meetings to go back in and pull these tweets out."* The group would query the dataset and read raw tweets together to understand the trends they were seeing in the charts. Reading and qualitative coding of individual messages was an important method used both in exploratory and focused analysis.

3.1.3. Analysis tools. Interviewees mentioned analyzing their data with a variety of tools. Most collected Twitter data using custom Python scripts and

Figure 1: Lariat comparing two groups ("victim death" vs. "victim survivor") and current search results ("rescue"). Highlighted are A) the keyword search box; B) the search result panel, with results for "rescue"; C) a list of saved groups; D) visualization of selected groups and search results, compared over the time dimension; E) tweets from a specific point on the visualization.

the Twitter API. One obtained tweets through the social data connector in DiscoverText. Once collected, data was stored in flat files (CSV and JSON), and sometimes in MySQL or MongoDB. Subsets of data were sometimes opened in Microsoft Excel or Google Spreadsheets, for qualitative analysis or simply for reading. Python was used to aggregate over the dataset and prepare quantitative summaries that were visualized in Tableau, Excel, or Google Spreadsheets. One interviewee reported creating semantic networks based on hashtag co-occurrence to visualize with Gephi. Another interviewee used computational modeling to analyze a large Twitter dataset: *"I have this little toolkit of common things that I do, like network diagrams, multi-dimensional scaling, time series, standard tools… and in combination, they tend to give a limited picture, but a pretty good picture of what's going on."* Combining tools and techniques gave our interviewees a more complete picture of the dataset.

3.1.4. Opportunities. Social scientists are asking many different kinds of questions about their data, but in the early phases of a project, they all struggle with EDA. Getting an overall sense of the content and dynamics in a large social media dataset can take months, and thus we saw a need for an easy-to-use exploratory visual analytics tool designed specifically for social media datasets.

3.2. Key features of Lariat

As a visual analytics tool for EDA, our two primary design goals of Lariat are to help users find a way to meaningfully analyze textual aspects of the dataset, and to enable comparisons across different segments of the data using quantitative visualizations. The system is intended to connect these two goals in an integrated, interactive fashion, so that users can switch between them easily and iterate on both aspects, accelerating the sensemaking process.

3.2.1. Building meaningful groupings. While researchers may often approach their datasets with some pre-existing ideas about what keywords and topics are significant, sometimes when analyzing a new dataset it is hard to know where to begin. To help users start thinking about the possible text based queries on the dataset, Lariat provides a list of keywords ordered by global frequency (Figure 2, Left).

Using words from this list, or their own words, users can search for tweets matching a search query (Figure 1, A). The results show up in the "Search Results" tab in the bottom-left panel (Figure 1, B). Users can read them and decide if they want to refine

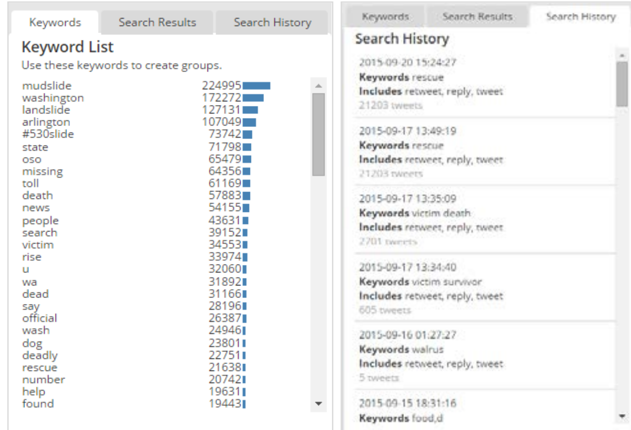

Figure 2: Additional tabs available in the bottom-left panel. Keywords (Left) shows keywords sorted by global frequency; Search History (Right) records past searches.

the search terms further. Currently the system supports simplified Boolean search query logic. All search terms are concatenated with "OR", but individual search terms can consist of multiple keywords concatenated by "AND". If a search term starts with NOT, the system will exclude tweets matching that term. For example, if a user wants to look at tweets related to "mudslide", but not the mudslide type of cake, the search terms could be "mudslide, NOT cake." If the user enters two keywords separated by a space, matching tweets must include both keywords (e.g., "victim survivor" and "victim death" in Figure 1, C). To facilitate the use of keywords, the search box provides auto-complete suggestions that pop up when users type in partial words.

When users find a search query that is potentially useful, the search terms can be saved as a group (Figure 1, C). Such saved groups can be revisited and edited later if desired. In addition to keyword-based search terms, users can filter out the tweets in the groups based on the tweets being originals, replies, or retweets.

3.2.2. Making visual comparisons. Saved groups, along with the results of the current search, can be used to create visualizations (Figure 1, D). Selected groups and search results are plotted in the visualization area using different colors, to help users make quantitative comparisons. At the moment the system supports visualizations that compare groups along nine different social media dimensions: 1) time, 2) time zones, 3) hashtags, 4) URLs, 5) whether the message contains a photo, 6) sentiment, 7) author name, 8) tweet type, and 9) mentions.

We arrived at this list of dimensions based on a systematic literature review of social media research papers and discussions with users in the initial testing stage. By selecting dimensions, users can examine the distribution of the data along different x-axes in the

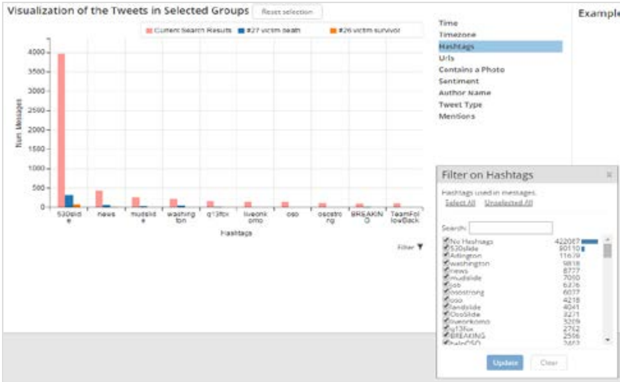

Figure 3: Visualizing the Hashtags dimension. For categorical dimensions like hashtags, the top 10 levels show on the chart. Users can use filters to view other levels.

visualization. The y-axis is always "number of tweets." Lariat has predefined rules to select the most appropriate type of visualization based on the type of dimension selected. For example, for a categorical dimension like *Hashtags*, a grouped bar chart is used (Figure 3), while selecting the *Time* dimension produces a time-series plot (Figure 1). Many of the categorical dimensions have very high cardinality (e.g. hundreds of thousands of *Author Names*). Therefore, for categorical dimensions, only the top 10 levels are shown on the visualization. Users can view other levels by applying filters to the dataset (Figure 3).

While visualizations can help users understand how their groups and search results compare, often researchers want to see raw data to better verify their assumptions and interpretations about the meaning of specific trends and patterns. Besides the Search Results, users can click on data points (the dots or bars in the visualizations) to see tweets that comprise each specific data point (Figure 1, E).

3.3.3. Sensemaking workflows. Sensemaking involves interlocked processes of creating schemas or frames and critically evaluating frames against data [21,30]. Lariat supports sensemaking by helping the user construct groups focused on the text-based content of messages, while also showing these groups in comparison to one another, and along several quantitative dimensions.

The sensemaking workflow in Lariat integrates categorical analysis of unstructured text-based messages together with visual analysis of quantitative aspects of the social media data. First, examination of tweet data in Lariat helps users build groups. Users can read raw tweets both in search results and by clicking on the visualization data points, which can generate ideas and refinements for queries and groups. The visualizations themselves can also lead to ideas for group refinement, if users notice a specific trend that goes with a certain level of the dimension (e.g. a hashtag or an account).

In the other direction, the construction of groups promotes more effective analysis. By building and refining groups, users have more meaningful and more precise analytical units for comparison and visualization. That is, building groups supports the comparison of numbers or trends in the visualization, and leads to more concrete hypotheses and directions for exploration.

4. Evaluation

We evaluated our final prototype to discover how it supports users in performing data exploration and how this support compares to their current practices. We aimed at naturalistic study focused on the specific application domain – exploration of Twitter data by social science researchers. Our evaluation followed the recommendations from a recent extensive review of evaluation practices of information visualization systems [22].

4.1. Research questions

We performed a qualitative study designed to answer the following research questions:

- How does Lariat support data exploration, and what role do different design elements play in this process?
- How does Lariat support discovery of knowledge and generation of insights?
- How does the functionality and user experience with Lariat compare to the other tools used by our participants?

While answering these research question we focused specifically on the six key functionalities we found crucial to support based on our formative research: 1) obtaining an overview of the dataset, 2) searching and navigating the dataset, 3) preserving intermediate exploration steps, 4) making comparisons, 5) tracking changes over time, 6) interacting naturally with tweet data.

4.2. Participants

We recruited 7 participants (3 female, 4 male) who have experience in Twitter data analysis. Participants' age varied from 20 to 50 years with 3 participants indicating age between 20-30 years. Two participants were undergraduate students; two were masters' students and one a PhD student. Finally, all the participants indicated that for data visualization and/or analysis they used MS Excel; four used Python or

Google Drive; three used Tableau; fewer used tools such as R, Gephi, Many Eyes, or MySQL.

4.3. Procedure

For the evaluation session Lariat was preloaded with a Twitter dataset concerning the March 2014 landslide near the city of Oso, Washington. This dataset was furnished by one of our interviewees in the formative study. For this study we filtered out non-English language tweets, leaving a total of 685,311 tweets. We ran participants through the study individually, with each session taking up to two hours. Each participant was rewarded with a $10 Amazon gift card.

The evaluation consisted of three parts. First, we introduced Lariat by describing its elements and demonstrating accomplishment of a short practice task. Then we asked our participants to perform exploration of the data on their own, but to make sure they were not lost we also proposed a task we drew from formative research that the participants were free to follow or disregard. The task asked the participants to explore the data trying to focus on different user accounts that were involved in dealing with the Oso landslide fallout. They were generally also encouraged to think aloud and explore the dataset independently of the task. After 30-40 minutes of exploration, we moved on to the next part.

The second part focused on collecting qualitative feedback on experiences with the system. Participants were asked 9 open questions regarding their experiences with the tool related to the 6 key functionalities as introduced earlier. For each question we also asked them which elements of the visualization helped them most, and we solicited their criticisms and suggestions. This component lasted around 30 minutes.

The final part was an open qualitative comparison of Lariat to other tools the participants have used before. They were asked 6 open questions focused on explicitly comparing Lariat to these other tools in relation to the key functionalities. They were encouraged to explain in detail their preferences, as well as describe in which situations in particular Lariat or the other tools would be more convenient for them.

During the study we ran Lariat on a MacBook Air with a 21.5" external monitor attached showing the tool in a Safari web browser. We video captured the screen and audio recorded each session. The tool itself was also instrumented with logging of user actions.

5. Findings

In order to quantify the interaction that the users had with our tool we logged their use of Lariat during the whole evaluation session. We coded our observer notes with a set of 7 categories: Ease of use, Functional limitations, Use of design elements, Functionality recommendations, Task/phase support, Insights, Usability issues. We discussed emerging high level themes by sequentially reviewing user comments in these 7 categories. In addition to the observation notes, we also analyzed logs of user interaction with Lariat. During the free exploration portion of the evaluation (30 minutes on average), users performed between 10 and 20 searches (mean 13). Participants generated about 48 visualizations each; these were rendered in response to search queries, selecting groups, and changing variables. The most used variables were Time (~28 times/person), Sender (~6 times/person), and Sentiment (~5 times/person). Users clicked on the visualizations between 6 and 61 times (mean 29 clicks). Participants created from 2 to 10 groups (mean 5). Groups had 1 to 3 keywords each. Most participants only refined their searches and groups once or twice, and rarely used exclusive keywords.

In our study, we observed users exploring the Oso dataset with Lariat. Participants moved back and forth between searching the dataset, reviewing search results, and visualizing the data. Sometimes reviewing visualizations or search results would lead participants to refine their search queries and continue, in an iterative sensemaking process. Lariat's flexibility to explore the data in different ways made it easy for participants to adapt to their needs. Based on our observation of users' free exploration with Lariat, we analyzed the insights that users uncovered about the dataset. We also asked participants to compare Lariat to other tools they used. Below, we discuss how participants used Lariat to understand the dataset, and the usefulness of Lariat's search, grouping, and visualization features.

5.1. Searching by keyword

The keyword search feature provided by Lariat helped participants uncover insight about the content in the dataset. Participants made observations about the types of tweets in the dataset by quickly trying out different searches and looking at the search results. One person discovered that a large number of tweets were from news channels, while another found that some of the most important groups involved in the event were underrepresented: *"Interesting! There were many search and rescue teams, but not many tweets about it" (P6)*.

From our formative studies, we knew that reading tweets is an important way in which the social scientists develop and validate insights, with links, photos, and account information all crucial metadata for interpreting tweet content. 5 out of 7 participants explicitly stated that they appreciated how Lariat surfaces rich contextual information from tweets: *"Super useful to have images in tweets" (P1)*. However, two of them indicated that this information also made it harder to quickly skim through large amounts of text. For some tasks, showing more tweets at one time in plain text would be preferable:

"Skimming is a bit harder [than in MySQL], because there are fewer tweets displayed and they have formatting. Reading many tweets is painful here"(P3)

Still 4 participants enjoyed the simplicity of the search feature, as it involves just entering keywords: *"I like it better [than MySQL], because there's no need to enter complex queries" (P2).* Others found it constrained, and felt they needed to build more complex queries to go deeper: *"It is good, but knowing how to code allows you to be more precise" (P3)*. All the participants uniformly agreed, however, that combination of search and immediate visualization was one of the best features of the tool: *"This is a really good tool, especially that you can visualize searches immediately" (P6).* While 5 of 7 users spent time mostly reading tweets, they also used the visualization directly to refine their searches, even without looking at the raw search results. For example, one user looked at a bar chart showing frequency of different accounts to inform the selection of new search keywords.

5.2. Visualizing groups

Saving and comparing groups made it easy to answer a variety of important questions about the dataset. One participant, while searching tweets relevant to the Oso landslide, began with the keyword "mudslide"; this matched almost half of the dataset, and he added exclusion keywords to refine the results. Later, he separately tried searching with "530slide", a more specific hashtag used during the event. He compared these two groups, and was able to see that "530slide" gave clearer results with less noise. This prompted him to reconsider starting with "mudslide" in the first place, and rethink his exploration strategy.

6 out of 7 participants explicitly stated that they liked the capability of creating groups of tweets by saving searches. They found that groups were useful for keeping track of exploration, making visualizations, and comparisons; they reported that tools they currently used did not provide any similar functionality. Still, two participants indicated that although helpful, building groups might not be that useful for their work. Different users also understood the functionality itself in different ways. Some of them saw groups as filters, others as saved visualizations, and yet others as "monitors" that dynamically capture tweets. Still, making visual comparisons between groups was a unique strength of Lariat:

"You have multiple queries and you can relate them to each other. This is something I can do here and I can't do with single queries elsewhere." (P1)

On the other hand, three participants explicitly stated that they felt a bit constrained by the types of comparisons they could do based on the queries they could build. One participant explained that there were different *layers* of comparison in their work, and Lariat only supported the lowest layer, closest to the data, but did not help with making higher-level inferences.

5.3. Temporal patterns

Exploration of temporal patterns in the dataset helped many participants make insightful observations of changes – dips and hills in the time-series visualization. When curious, some participants clicked on anomalies in the chart to view related tweets in detail. The time-series visualizations also helped users see the sequence of events, starting from the onset of the landslide itself. Such sequential, time-ordered exploration helped one user make a surprising discovery:

"Almost 2 days after the actual event - there was a boy found - it was interesting to find that" (P2)

Time-series charts also helped some users make insightful comparisons between groups and searches. One participant visualized the volume of tweets from 3 groups he had created earlier: "victims", "rescue" and "survivors". Comparing these on the time-series chart led him to observe clear differences in the changing impact of the groups over time. Specifically, he noticed that survivors and rescues were positively correlated in time, but that the survivors and victims seemed inversely correlated.

All participants found the ability to visualize the data over time to be important. They felt that it was easier to create time-series visualizations, and to make time-related comparisons in Lariat than with many other tools: "*It is on par with Tableau, it is better in terms of ease and efficiently than R, Python" (P7).* Two participants, however, felt that they would have more detailed control using SQL queries or programming tools such as R or Python: *"Right now, I can do less, in terms of time, than in a SQL query" (P1).*

5.4. Social media metadata

Metadata was also a source of insightful information for our participants. The use of time zone as a proxy for geographical location triggered some users to investigate the global reach of information about the disaster. One user found tweets about the landslide in places as distant from the event as Pretoria, South Africa:

"Interesting to see how far a particular hashtag went. [...] That is really interesting that there are tweets from Pretoria that are still related. I would be curious to explore that a bit more." (P7)

For 2 users, it was surprising, and even confusing, to be able to find tweets about what they had considered a local event being exchanged around the world. Other useful metadata included the links embedded in tweets and user accounts. Being able to easily see such information along with the tweet prompted some users to explore more deeply, by following links and visiting profiles on Twitter.

Lariat also provides sentiment analysis as a dimension for visualizing the dataset through an off-the-shelf python package [1]. Although, as we mentioned before, it is very imprecise, this feature was explicitly mentioned as helpful by 3 of 7 participants. Specifically, one user, exploring the keyword "landslide", used the sentiment visualization to find examples of *positive* tweets; she was interested in what could possibly be positive about a disaster event. This led her to discover that landslide was sometimes used positively to indicate a "landslide victory" in a local election, unrelated to the disaster. She refined the search to exclude such noise. As in this example, users had existing hypotheses about the dataset, which they evaluated by inspecting tweets. Users enjoyed visualizing sentiment because of their intuitive expectations of what kind of positive, negative or neutral tweets they could expect. One user expected that neutral tweets would represent dry, emotionless reporting of information, such as from news media sources. After using the sentiment visualization to explore the neutral tweets, she verified this expectation:

"I would imagine that neutral is news reporting, and the people would be more emotional. It's good to know that based on sentiment more tweets are from news." (P7)

Tweets are socially generated information and the source of this information is often of particular interest to social science researchers. Our participants were interested in learning more about user accounts. Lariat displays tweets in a format similar to Twitter, containing the user name and profile picture along with the tweet contents. One of our participants became interested in tweets created by a specific user:

"Oh, I found a researcher. I would like to know who it is… he has a PhD, but also reports in real time. I would like to know more about who that person is." (P7)

6 out of 7 users found the visualizations of social media dimensions provided in Lariat straightforward and simple, compared to the other tools they used, because they were specifically oriented towards social media data: *"This is more straightforward. It is designed for tweets. Tableau is not" (P6).* However, some also felt that the visualization functionality was already covered by other commercial tools, and that the implementation in Lariat was limited in some ways. Participants sometimes wanted to relate multiple dimensions at once, or to visualize other kinds of metadata that were not included in Lariat.

6. Discussion

Below, we discuss several implications for design based on our findings from evaluating the Lariat tool itself as well as the design process we followed.

6.1. Integration with tweets

Almost all participants appreciated the way that Lariat displayed the contents of tweets in combination with rich contextual information not provided by other tools. Nevertheless, users wanted to be able to more easily interact with the contents of the tweets within the tool. For example, they found it tedious to copy text into the search box from the tweets in order to refine their searches. They also wanted to be able to directly click on links or user profiles to follow them up, either in Lariat itself, or even outside of it. Another observation relates to the presentation of the tweets themselves: while preserving the "Twitter look" was appreciated, participants also wanted to be able to skim a larger amount of tweets for some tasks, and to see more tweets at any one time. Such dualism of opinions, expressed sometimes by one and the same user, suggests that the form of display of tweets should be adjustable, to better support different tasks.

6.2. Annotating the data

Four participants explicitly stated that they used the saved groups to keep track of their exploration, and said that having a search history made them feel safe about being able to navigate their exploration process easily. There were, however, a number of comments related to the support that Lariat could offer for analysis of events. Specifically, users indicated annotating parts of the dataset, either related to time or contents, or specifically labeling some of the interesting findings, could be useful. Therefore to move from exploration into more focused analysis,

[1] TextBlob (https://textblob.readthedocs.io/)

tools should provide more sophisticated and flexible ways of tracking and annotating data.

6.3. The analysis ecosystem

While users found it easy to explore Twitter data in Lariat, they also felt that they would need to go outside of the tool to do more complex analysis. 4 of 7 users explicitly stated that they wanted to export results from Lariat to do further analysis: *"This is good for exploring the dataset itself, for initial search […] later when we need deeper information it would not help that much" (P6)*. On the input side, one user indicated that, for his work, it would be useful to load a dataset into Lariat with custom labeling already applied. While tools such as Lariat could support analytical qualitative coding, it would make the tool more complex, possibly negatively affecting its simplicity and ease of use. Alternatively, perhaps designers should look for opportunities to integrate such tools with the ecosystem of tools and technologies that social scientists currently use. Lariat could import datasets in different formats and support further analysis in other tools by providing an easy export function. This balance between ease of use and breadth/depth of functionality is an important consideration for similar tools.

6.4. Multi-level analysis and mini-schemas

As described in the results section, all the users appreciated the ability to build categories to structure their exploration of search queries. Still, our participants did not see Lariat as a tool for making higher-level inferences. One of the users explicitly indicated that different levels of analysis might require different kinds of support from the tool. Lariat's search query-based groups were perceived as very low-level and close to the data; while useful, this does not directly produce higher-level findings: *"This is not journaling article level, trying to describe interactions between things is hard here."* (P1). As the original design of the grouping function aimed to support schematizing actions in Pirolli and Card's sensemaking notional model, we found the model could be extended to include "mini-schemas" into the loop. That is, the low-level task of creating groups is a way to generate small and clean mini-schemas, and these mini-schemas can be used to construct higher-level schemas or serve as a probe to collect and search for more information. We want to argue that this point is specifically useful for social media data because of the infinite boundary of related data beyond the tweet itself. For instance, by searching and creating groups, one may find official Twitter accounts for news channels play a significant role in broadcasting the information of the disaster. This "mini-schema" of new channel accounts can be used to seek more details, such as their subscribers and the subscribers' behaviors during the disaster event. These mini-schemas may not provide a full story yet, but they can be valuable units that one may leverage for higher-level analyses. Additionally, these low-level mini-schemas are cleaner and may be easier to generate meaningful results through automatic methods because there will be less noise in the input. We leave the completion of extending the sensemaking theory with the mini-schemas and applying more computational methods for future work.

7. Future work and conclusion

In this paper, we presented Lariat, a tool for exploring social media datasets through integrated search, visualization, and category building, and the results of a formative study of social media researchers studying Twitter. We evaluated Lariat in an exploratory qualitative study; participants were able to use Lariat to obtain insights about various aspects of their datasets and found its simple search and visualization features to be particularly useful. It is evident from our formative study and the evaluation of Lariat, that visual analytics tools in this domain must allow deeper interaction with social media data, need to fit into the ecosystem of data analysis tools, and must be designed to support analysis on multiple levels of abstraction.

For future work, we will explore the ways to add automatic methods based on the groups, aiming to support higher-level analyses. We also want to test Lariat with a broader range of social media researchers. As social media research is thriving and the field of visual analytics continues to build an understanding of different user communities, these findings can support researchers in designing better exploratory analysis tools for social media researchers.

8. Acknowledgments

We would like to thank everyone who participated in the formative studies and the evaluation, and the reviewers whose comments greatly improved the paper.

9. References

[1] Socialpeeks. Retrieved from http://socialpeeks.com/

[2] Boyd, D. Why youth (heart) social network sites: The role of networked publics in teenage social life. *MacArthur foundation series on digital learning–Youth, identity, and digital media volume* (2007), 119-142.

[3] Brooks, M., Robinson, J. J., Torkildson, M. K. and Aragon, C. R. Collaborative Visual Analysis of Sentiment in Twitter Events. In *Cooperative Design, Visualization, and Engineering*, (2014), 1-8.

[4] Carpendale, S. Evaluating information visualizations. In *Information Visualization*, (2008), 19-45.

[5] Chin Jr, G., Kuchar, O. A. and Wolf, K. E. Exploring the analytical processes of intelligence analysts. In *Proc. of CHI*, ACM (2009), 11-20.

[6] Dörk, M., Gruen, D., Williamson, C. and Carpendale, S. A visual backchannel for large-scale events. *IEEE Transactions on Visualization and Computer Graphics, 16*, 6 (2010), 1129-1138.

[7] Diakopoulos, N., Naaman, M. and Kivran-Swaine, F. Diamonds in the rough: Social media visual analytics for journalistic inquiry. In *Proc. Visual Analytics Science and Technology (VAST),* IEEE (2010), 115-122.

[8] Diesner, J., Frantz, T. L. and Carley, K. M. Communication networks from the Enron email corpus "It's always about the people. Enron is no different". *Computational & Mathematical Organization Theory*, *11*, 3 (2005), 201-228.

[9] Goggins, S. P., Laffey, J. and Gallagher, M. Completely online group formation and development: small groups as socio-technical systems. *Information Technology & People*, *24*, 2 (2011), 104-133.

[10] Goggins, S. P., Mascaro, C. and Valetto, G. Group informatics: A methodological approach and ontology for sociotechnical group research. *Journal of the American Society for Information Science and Technology*, *64,* 3 (2013), 516-539.

[11] Hubmann-Haidvogel, A., Brasoveanu, A. M., Scharl, A., Sabou, M. and Gindl, S. Visualizing contextual and dynamic features of micropost streams (2012).

[12] Kandel, S., Paepcke, A., Hellerstein, J. M. and Heer, J. Enterprise data analysis and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics*, *18*, 12 (2012), 2917-2926.

[13] Kang, Y.-a. and Stasko, J. Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study. In *Proc. of VAST*, IEEE (2011), 21-30.

[14] Keim, D. A., Kohlhammer, J., Ellis, G. and Mansmann, F. *Mastering the information age-solving problems with visual analytics*. Florian Mansmann, 2010.

[15] Klein, G., Moon, B. and Hoffman, R. R. Making sense of sensemaking 2: A macrocognitive model. *Intelligent Systems, IEEE*, *21*, 5 (2006), 88-92.

[16] Maddock, J., Starbird, K., Al-Hassani, H. J., Sandoval, D. E., Orand, M. and Mason, R. M. Characterizing online rumoring behavior using multi-dimensional signatures. In *Proc. of CSCW*, ACM (2015), 228-241.

[17] Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S. and Miller, R. C. Twitinfo: aggregating and visualizing microblogs for event exploration. In *Proc. of CHI*, ACM (2011), 227-236.

[18] Mazumdar, S., Ciravegna, F., Gentile, A. L. and Lanfranchi, V. Visualising context and hierarchy in social media. In *Proc. International Workshop on Intelligent Exploration of Semantic Data (IESD'2012) at EKAW* (2012).

[19] Munzner, T. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, *15*, 6 (2009), 921-928.

[20] Perer, A. and Shneiderman, B. Integrating statistics and visualization for exploratory power: From long-term case studies to design guidelines. *IEEE Computer Graphics and Applications*, 3 (2009), 39-51.

[21] Pirolli, P. and Card, S. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proc. of international conference on intelligence analysis* (2005), 2-4.

[22] Plaisant, C., Fekete, J.-D. and Grinstein, G. Promoting insight-based evaluation of visualizations: From contest to benchmark repository. *IEEE Transactions on Visualization and Computer Graphics*, *14*, 1 (2008), 120-134.

[23] Russell, D. M., Stefik, M. J., Pirolli, P. and Card, S. K. The cost structure of sensemaking. In *Proc. of the INTERACT'93 and CHI'93,* ACM (1993), 269-276.

[24] Sedlmair, M., Isenberg, P., Baur, D., Mauerer, M., Pigorsch, C. and Butz, A. Cardiogram: visual analytics for automotive engineers. In *Proc. of CHI*, ACM (2011), 1727-1736.

[25] Shulman, S. DiscoverText: Software training to unlock the power of text. In *Proc. of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, ACM (2011), 373-373.

[26] Sloan, L., Morgan, J., Housley, W., Williams, M., Edwards, A., Burnap, P. and Rana, O. Knowing the tweeters: Deriving sociologically relevant demographics from Twitter. *Sociological research online*, *18*, 3 (2013), 7.

[27] Starbird, K. and Palen, L. (How) will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. In *Proc. of CSCW* , ACM (2012), 7-16.

[28] Tory, M. and Staub-French, S. Qualitative analysis of visualization: A building design field study. In *Proc. of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, ACM (2008), 7.

[29] Tukey, J. W. *Exploratory Data Analysis*. Addison-Wesley Publishing Company, 1977.

[30] Zagheni, E., Garimella, V. R. K. and Weber, I. Inferring international and internal migration patterns from twitter data. In *Proc. of the companion publication of the 23rd international conference on World wide web companion*, (2014), 439-444.